STIC Biotechnology Systems Branch

RAW SEQUENCE LISTING
ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: _10/553,676_
Source: _1FWP_
Date Processed by STIC: _8/15/06_

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.
PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:
1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY
FOR CRF SUBMISSION AND PATENTIN SOFTWARE QUESTIONS, PLEASE CONTACT
MARK SPENCER, TELEPHONE: 571-272-2510; FAX: 571-273-0221

TO **REDUCE** ERRORED SEQUENCE LISTINGS, **PLEASE** USE THE **CHECKER VERSION 4.4.0 PROGRAM**, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:
http://www.uspto.gov/web/offices/pac/checker/chkrnote.htm

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.
Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.
Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (**<http://www.uspto.gov/ebc/efs/downloads/documents.htm>** , **EFS Submission User Manual** - ePAVE)
2. **U.S. Postal Service:** Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450
3. **Hand Carry, Federal Express, United Parcel Service, or other delivery service (EFFECTIVE 01/14/05):**
   U.S. Patent and Trademark Office, Mail Stop Sequence, Customer Window, Randolph Building, 401 Dulany Street, Alexandria, VA 22314

Revised 01/10/06

Raw Sequence Listing Error Summary

ERROR DETECTED	SUGGESTED CORRECTION	SERIAL NUMBER: $10/553,676$

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

1____Wrapped Nucleics
Wrapped Aminos
The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor **after** creating it. Please adjust your right margin to .3; this will prevent "wrapping."

2____Invalid Line Length
The rules require that a line **not exceed 72** characters in length. This includes white spaces.

3_✓_Misaligned Amino
Numbering
The numbering under each 5th amino acid is misaligned. Do **not** use tab codes between numbers; use **space characters**, instead.

4_✓_Non-ASCII
The submitted file was **not** saved in ASCII(DOS) text, as **required** by the Sequence Rules. **Please ensure your subsequent submission is saved in ASCII text.**

5____Variable Length
Sequence(s)____contain n's or Xaa's representing more than one residue. **Per Sequence Rules, each n or Xaa can only represent a single residue.** Please present the **maximum** number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.

6____PatentIn 2.0
"bug"
A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) _____. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. **This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.**

7____Skipped Sequences
(OLD RULES)
Sequence(s) _____ missing. If intentional, please insert the following lines for **each** skipped sequence:
(2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
(i)   SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)
(xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
This sequence is intentionally skipped
Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to **include** the skipped sequences.

8____Skipped Sequences
(NEW RULES)
Sequence(s) _____ missing. If **intentional**, please insert the following lines for **each** skipped sequence.
<210> sequence id number
<400> sequence id number
000

9____Use of n's or Xaa's
(NEW RULES)
Use of n's and/or Xaa's have been detected in the Sequence Listing.
Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.
In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.

10____Invalid <213>
Response
Per 1.823 of Sequence Rules, the only **valid** <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is **required** when <213> response is Unknown or is Artificial Sequence. (see item 11 below)

11____Use of <220>
Sequence(s) ___ missing the <220> "Feature" and associated numeric identifiers and responses. Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section or use "chemically synthesized" as explanation. (See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32), also Sec. 1.823 of Sequence Rules

12____PatentIn 2.0
"bug"
Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.

13____ Misuse of n/Xaa
"n" can only represent a single <u>nucleotide</u>; "Xaa" can only represent a single <u>amino acid</u>

RAW SEQUENCE LISTING   DATE: 08/15/2006
PATENT APPLICATION: US/10/553,676   TIME: 12:21:39

            Input Set : A:\PTO.RJ.txt
            Output Set: N:\CRF4\08152006\J553676.raw

see item 4 on Error Summary Sheet

```
    5 <110> APPLICANT: MERCK-SANTE
    6        CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE (CNRS)
    9 <120> TITLE OF INVENTION: Insulin-induced gene as therapeutic target in diabetes
   13 <130> FILE REFERENCE: BFF 03P0004
C--> 17 <140> CURRENT APPLICATION NUMBER: US/10/553,676
C--> 17 <141> CURRENT FILING DATE: 2005-10-17
   17 <160> NUMBER OF SEQ ID NOS: 8
   21 <170> SOFTWARE: PatentIn version 3.1
```

Suggestion: consult Sequence Rules for VALID format

Does Not Comply Corrected Diskette Needed

see pp 1-4

ERRORED SEQUENCES

```
   73 <210> SEQ ID NO: 2
   75 <211> LENGTH: 353
   77 <212> TYPE: PRT
   79 <213> ORGANISM: Rattus sp.
   83 <400> SEQUENCE: 2
   85 Met Leu Cys Thr Leu Phe Leu Leu Leu Leu Ala Leu Gly Ile Val Gln
E--> 86 1         5         10             15
   89 Thr Thr Arg Pro Cys Phe Pro Gly Cys Gln Cys Glu Glu Glu Thr Phe
E--> 90           20    25         30               30
   93 Gly Leu Phe Asp Ser Phe Ser Leu Ile Arg Val Asp Cys Ser Ser Leu
E--> 94       35        40        45
   97 Gly Pro His Ile Val Pro Val Pro Ile Pro Leu Asp Thr Ala His Leu
E--> 98    50        55        60
  101 Asp Leu Ser Ser Asn Arg Leu Glu Thr Val Asn Glu Ser Val Leu Gly
E--> 102 65        70        75        80
  105 Gly Pro Gly Tyr Thr Thr Leu Ala Gly Leu Asp Leu Ser His Asn Leu
E--> 106       85        90        95
  109 Leu Thr Ser Ile Thr Pro Thr Ala Phe Ser Arg Leu Arg Tyr Leu Glu
E--> 110    100       105       110
  113 Ser Leu Asp Leu Ser His Asn Gly Leu Ala Ala Leu Pro Ala Glu Val
E--> 114    115       120       125
  117 Phe Thr Ser Ser Pro Leu Ser Asp Ile Asn Leu Ser His Asn Arg Leu
E--> 118 130       135       140
  121 Arg Glu Val Ser Ile Ser Ala Phe Thr Thr His Ser Gln Gly Arg Ala
E--> 122 145       150       155       160
  125 Leu His Val Asp Leu Ser His Asn Leu Ile His Arg Leu Leu Pro Tyr
E--> 126       165       170       175
  127 Pro Ala Arg Ala Ser Leu Ser Ala Pro Thr Ile Gln Ser Leu Asn Leu
E--> 128    180       185       190
  131 Ser Trp Asn Arg Leu Arg Ala Val Pro Asp Leu Arg Asp Leu Pro Leu
E--> 132 195       200       205
```

misaligned amino acid numbers. see item 3 on Error Summary Sheet

see P.2

RAW SEQUENCE LISTING   DATE: 08/15/2006
PATENT APPLICATION: US/10/553,676   TIME: 12:21:39

Input Set : A:\PTO.RJ.txt
Output Set: N:\CRF4\08152006\J553676.raw

```
      135 Arg Tyr Leu Ser Leu Asp Gly Asn Pro Leu Ala Thr Ile Asn Pro Gly
E-->  136     210         215         220
      139 Ala Phe Met Gly Leu Ala Gly Leu Thr His Leu Ser Leu Ala Ser Leu
E-->  140 225         230         235         240
      143 Gln Gly Ile Leu Gln Leu Pro Pro His Gly Phe Arg Glu Leu Pro Gly
E-->  144         245         250         255
      147 Leu Gln Val Leu Asp Leu Ser Gly Asn Pro Lys Leu Lys Trp Ala Gly
E-->  148         260         265         270
      151 Ala Glu Val Phe Ser Gly Leu Gly Leu Leu Gln Glu Leu Asp Leu Ser
E-->  152     275         280         285
      155 Gly Ser Ser Leu Val Pro Leu Pro Glu Thr Leu Leu His His Leu Pro
E-->  156     290         295         300
      159 Ala Leu Gln Ser Val Ser Val Gly Gln Asp Val Gln Cys Arg Arg Leu
E-->  160 305         310         315         320
      163 Val Arg Glu Gly Ala Val His Arg Gln Pro Gly Ser Ser Pro Lys Val
E-->  164         325         330         335
      167 Val Leu His Cys Gly Asp Thr Gln Glu Ser Ala Arg Gly Pro Asp Ile
E-->  168         340         345         350
      171 Leu
      339 <210> SEQ ID NO: 4
      341 <211> LENGTH: 353
      343 <212> TYPE: PRT
      345 <213> ORGANISM: Homo sapiens
      349 <220> FEATURE:
      351 <221> NAME/KEY: misc_feature
      353 <222> LOCATION: (121)..(121)
      355 <223> OTHER INFORMATION: 'Xaa' in position 121 represents Ala or Thr.
      357 <400> SEQUENCE: 4
      359 Met Pro Trp Pro Leu Leu Leu Leu Leu Ala Val Ser Gly Ala Gln Thr
E-->  360 1           5           10          15
      363 Thr Arg Pro Cys Phe Pro Gly Cys Gln Cys Glu Val Glu Thr Phe Gly
E-->  364         20          25          30
      367 Leu Phe Asp Ser Phe Ser Leu Thr Arg Val Asp Cys Ser Gly Leu Gly
E-->  368     35          40          45
      371 Pro His Ile Met Pro Val Pro Ile Pro Leu Asp Thr Ala His Leu Asp
E-->  372     50          55          60
      375 Leu Ser Ser Asn Arg Leu Glu Met Val Asn Glu Ser Val Leu Ala Gly
E-->  376 65          70          75          80
      379 Pro Gly Tyr Thr Thr Leu Ala Gly Leu Asp Leu Ser His Asn Leu Leu
E-->  380     85          90          95
      383 Thr Ser Ile Ser Pro Thr Ala Phe Ser Arg Leu Arg Tyr Leu Glu Ser
E-->  384     100         105         110
W-->  387 Leu Asp Leu Ser His Asn Gly Leu Xaa Ala Leu Pro Ala Glu Ser Phe
E-->  388     115         120         125
      391 Thr Ser Ser Pro Leu Ser Asp Val Asn Leu Ser His Asn Gln Leu Arg
E-->  392     130         135         140
      395 Glu Val Ser Val Ser Ala Phe Thr Thr His Ser Gln Gly Arg Ala Leu
E-->  396 145         150         155         160
      399 His Val Asp Leu Ser His Asn Leu Ile His Arg Leu Val Pro His Pro
```

same error (handwritten annotation)

misaligned number (handwritten annotation)

RAW SEQUENCE LISTING   DATE: 08/15/2006
PATENT APPLICATION: US/10/553,676   TIME: 12:21:39

Input Set : **A:\PTO.RJ.txt**
Output Set: **N:\CRF4\08152006\J553676.raw**

```
E--> 400          165          170          175
     403 Thr Arg Ala Gly Leu Pro Ala Pro Thr Ile Gln Ser Leu Asn Leu Ala
E--> 404          180          185          190
     407 Trp Asn Arg Leu His Ala Val Pro Asn Leu Arg Asp Leu Pro Leu Arg
E--> 408          195          200          205
     411 Tyr Leu Ser Leu Asp Gly Asn Pro Leu Ala Val Ile Gly Pro Gly Ala
E--> 412     210          215          220
     415 Phe Ala Gly Leu Gly Gly Leu Thr His Leu Ser Leu Ala Ser Leu Gln
E--> 416 225          230          235          240
     419 Arg Leu Pro Glu Leu Ala Pro Ser Gly Phe Arg Glu Leu Pro Gly Leu
E--> 420          245          250          255
     423 Gln Val Leu Asp Leu Ser Gly Asn Pro Lys Leu Asn Trp Ala Gly Ala
E--> 424          260          265          270
     427 Glu Val Phe Ser Gly Leu Ser Ser Leu Gln Glu Leu Asp Leu Ser Gly
E--> 428          275          280          285
     431 Thr Asn Leu Val Pro Leu Pro Glu Ala Leu Leu Leu His Leu Pro Ala
E--> 432     290          295          300
     435 Leu Gln Ser Val Ser Val Gly Gln Asp Val Arg Cys Arg Arg Leu Val
E--> 436 305          310          315          320
     439 Arg Glu Gly Thr Tyr Pro Arg Arg Pro Gly Ser Ser Pro Lys Val Ala
E--> 440          325          330          335
     442 Leu His Cys Val Asp Thr Arg Glu Ser Ala Ala Arg Gly Pro Thr Ile
E--> 443          340          345          350
     446 Leu
     491 <210> SEQ ID NO: 8
     493 <211> LENGTH: 29
     495 <212> TYPE: DNA
C--> 497 <213> ORGANISM: Artificial () primer
W--> 501 <220> FEATURE:
W--> 501 <223> OTHER INFORMATION:
W--> 501 <400> 8
     502 gatggaaaga gctcttacat gtgtttatt                    29
E--> 504 1
E--> 506 P03139 VE
W--> 508 1
```

[Handwritten annotations:]

same erros

delete

insert <2207 (no response. <2207 is a "header" only)

move "primer" to <2237 line

<2237 primer (the explanation for <2137 Artificial Sequence goes in <2207-<2237 section, not on <2137 line)

delete

This error also appears in Sequences 5 through 7.

```
<400> 3
tccagccccc acc atg ccg tgg ccc ctg ctg ctg ctg ctg gcc gtg agt        49
         Met Pro Trp Pro Leu Leu Leu Leu Leu Ala Val Ser ———→
         1         5       5   10                      10
```

(misaligned numbers)

Move amino acids directly

under their codons

the above is a sample of global

errors in Sequence 3

VERIFICATION SUMMARY   DATE: 08/15/2006
PATENT APPLICATION: **US/10/553,676**   TIME: 12:21:40

Input Set : **A:\PTO.RJ.txt**
Output Set: **N:\CRF4\08152006\J553676.raw**

```
L:17 M:270 C: Current Application Number differs, Replaced Current Application No
L:17 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:86 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:2
M:332 Repeated in SeqNo=2
L:194 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:3,Line#:190
L:197 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:201 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:205 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:209 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:213 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:217 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:221 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:224 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 after pos.:385
L:225 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:229 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:233 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:237 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:241 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:245 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:249 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:255 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:259 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:263 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:267 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:271 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:275 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:279 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:283 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:287 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:360 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:4
M:332 Repeated in SeqNo=4
L:387 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:112
L:456 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:5
L:459 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ#:5, <213>
ORGANISM:Artificial Sequence
L:459 M:258 W: Mandatory Feature missing, <223> Tag not found for SEQ#:5, <213>
ORGANISM:Artificial Sequence
L:459 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:5,Line#:459
L:469 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:6
L:473 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ#:6, <213>
ORGANISM:Artificial Sequence
L:473 M:258 W: Mandatory Feature missing, <223> Tag not found for SEQ#:6, <213>
ORGANISM:Artificial Sequence
L:473 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:6,Line#:473
L:483 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:7
L:487 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ#:7, <213>
ORGANISM:Artificial Sequence
L:487 M:258 W: Mandatory Feature missing, <223> Tag not found for SEQ#:7, <213>
ORGANISM:Artificial Sequence
L:487 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:7,Line#:487
L:497 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:8
L:501 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ#:8, <213>
ORGANISM:Artificial Sequence
```

L:501 M:258 W: Mandatory Feature missing, <223> Tag not found for SEQ#:8, <213> ORGANISM:Artificial Sequence
L:501 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:8,Line#:501

VERIFICATION SUMMARY   DATE: 08/15/2006
PATENT APPLICATION: **US/10/553,676**   TIME: 12:21:40

Input Set : **A:\PTO.RJ.txt**
Output Set: **N:\CRF4\08152006\J553676.raw**

L:504 M:254 E: No. of Bases conflict, this line has no nucleotides.
L:506 M:334 E: (2) Invalid Amino Acid in Coding Region, NUMBER OF INVALID KEYS:2
L:508 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:8